Response to gmd-2016-197-RC1:

2. Page 2 line 20: “This is because different computing architectures are less or more suitable to increasing problem size along any of these axes.” My reading did not allow me to understand this sentence concluding the paragraph. Maybe a little more explanation would help (me, at least)

   See page 2, line 20. We agree that this line is a bit obscure, and have added a paragraph of explanation and examples.

3. Page 3 line 26: “One issue was the limitations imposed by memory.” I suggest replacing memory by memory bandwith (e.g. not memory size)

   Agreed, see page 4, line 3.

4. Page 4 and 5: I’m not fully convinced by the Figure 1: is seems incomplete (and difficult to complete!). One could add a number of boxes (OcnDyn, OcnPhy, IceDyn, IceBio, etc...) so as a number of couplings, interactions and related processes. Would an “incomplete list” be better, or is there a better way to build a picture to make this complex ESM architecture visible?

   We have added additional text to the caption to Fig. 1 to make it clear that the tree shown in this diagram could have further embedded components.

5. Page 5 line 8: “but a model at fixed resolution is capped in terms of time to solution, absent advances in hardware or algorithm.” I agree if you are still talking about one dycore in this paragraph, but this is not true for a whole ESM.

   Our understanding of this comment is that adding complexity can improve scaling, but in fact it cannot improve time to solution. We have added some text to explain this, and connect to the next paragraph, which goes in depth into the issue of model complexity. See page 6, line 1.

6. Page 6 Figure 2: Not clear to me what is the conclusion of paragraph ending line 14, nor precisely what is demonstrated on Figure 2.

   This is a fair comment, as the discussion around Fig. 2 foreshadows a discussion to come, further down in Sec. 3.3. We have added some lines to make the link to the discussion of coupling cost more explicit, which we hope addresses the concern raised here. See page 6, line 23. We have also modified Fig. 2 to show the bounding rectangle.

7. Page 8: I believe the answers to the questions listed at top of page depend on how the computer platform is used: in dedicated mode, or not? But this seems to be taken in account in the ASYPD metric?

   Yes, this is correct. The ASYPD metric is supposed to reveal performance issues associated with such “policy” issues such as whether the target machine is in “dedicated mode” as the reviewer has stated. In particular, we wish to see ASYPD measured just as it is run in production, not under ideal conditions: so one should not take the measurement in dedicated mode, if in practice the machine is not dedicated! See also page 8, line 1; page 11, line 4; page 13, line 2; page 16, line 6; page 20, line 21; page 23, line 1.

8. Page 11: taking the resolution in account Here, you have only taken the spatial resolution in account. I wonder if somehow (not simple though), temporal resolution should be take in account in the G metric: using a large or a small time step does indeed widely change the number of floating point operation you need for one simulated year. Now, what time step to use (those are different in atmosphere, ocean, coupler...)? I do not have any simple answer, but to be able to compare metrics between ESMS, I suggest to add something in G to take in account number of time steps per year as an unavoidable constraint on the number of floating point operations. (Could also be in a next generation of those metrics)

   We do not believe the temporal resolution should be included in G, for the following reason: it is chosen as a consequence of the spatial resolution and the accuracy one needs to achieve. For example, the timestep may be bounded by a CFL criterion once the spatial resolution is given, for one class of methods. For implicit methods where there is no CFL, it is still bounded by the level of accuracy needed. In any case, the timestep is never varied independently of the limits of physics and resolution purely for reasons of speed and performance, and is generally set to the largest value imposed by those limits. It therefore should not be separately accounted for in G.

9. Page 19: Reference for XIOS: is this paper from Joussaume et al. the appropriate reference for XIOS today?

   Unfortunately, there does not yet exist a better reference to XIOS than Joussaume et al. (2012) and the URL to the source and documentation, already provided.